文章编号：0454-6296（2000）03-0233-09

棉铃虫核型多角体病毒几丁质酶基因及杆状病毒几丁质酶基因的分子进化

张传溪[1]，林欣大[1]，吴 峻[2]

（1. 浙江大学应用昆虫学研究所，杭州 310029；
2. 中国科学院上海生物化学研究所，上海 200031）

摘要：用PCR方法扩增了棉铃虫 *Helicoverpa armigera* 单粒包埋型核型多角体病毒（HaSNPV）几丁质酶基因，测定了基因编码区的核苷酸全序列。基因编码区全长1 713 bp，可编码570个氨基酸残基组成的多肽，预计分子量为63.6 kD。将所推导的HaSNPV几丁质酶氨基酸序列与其它已知杆状病毒几丁质酶氨基酸序列进行联配比较，结果表明HaSNPV与谷实夜蛾 *H. zea* 单粒包埋型核型多角体病毒（HzSNPV）的氨基酸序列非常相似，同源性高达90.7%，与苜蓿丫纹夜蛾 *Autographa californica* 多粒包埋型核型多角体病毒（AcMNPV）、家蚕 *Bombyx mori* 核型多角体病毒（BmNPV）、美国白蛾 *Hyphantria cunea* 核型多角体病毒（HcNPV）、舞毒蛾 *Lymantria dispar* 多粒包埋型核型多角体病毒（LdMNPV）、黄杉毒蛾 *Orgyia pseudotsugata* 多粒包埋型核型多角体病毒（OpMNPV）和云杉卷叶蛾 *Choristoneura fumiferana* 核型多角体病毒（CfMNPV）氨基酸序列同源性分别为64.4%、64.9%、64.2%、62.9%、66.2%和61.5%。根据氨基酸序列用PC\GENE程序绘制已知杆状病毒几丁质酶的分子系统树，并与杆状病毒中最为保守的多角体蛋白基因系统树作了比较，结果表明几丁质酶基因和多角体蛋白基因的进化速率是不尽相同的。

关键词：棉铃虫核型多角体病毒；几丁质酶基因；核苷酸序列；杆状病毒；分子系统树
中图分类号：S476.13；Q751 **文献标识码**：A

棉铃虫 *Helicoverpa armigera* 是棉花、玉米、烟草、番茄、小麦等经济作物的大害虫。化学农药的大量使用，使生态平衡遭受破坏，棉铃虫抗药性剧增。近年棉铃虫在我国广大棉区猖獗成灾，造成减产甚至绝收。棉铃虫核型多角体病毒（HaNPV）是我国第一个商品病毒杀虫剂，具使用安全，害虫不产生抗性等优点，是一种很有发展潜力的可替代生物农药，但HaNPV与其它病毒杀虫剂一样亦存在杀虫致死速度较慢等缺点。因此，通过遗传工程构建毒力强，即效性高的重组HaNPV具有重要的意义。对HaNPV进行分子生物学研究，是今后进行其重组研究的必不可少的基础。昆虫病毒多角体被宿主昆虫吞食后，在病毒粒子进入中肠细胞之前，需要穿过含几丁质的中肠围食膜。感病昆虫死亡后，含几丁质的昆虫体壁迅速酶解，释放出体内的病毒多角体，感染其它昆虫个体。这一酶解过程在确保病毒能通过风和雨溅等自然力量进行有效扩散中可能起着重要的作用。已知NPV编码的几丁质酶在体壁酶解过

基金项目：国家自然科学基金资助项目（39770031）和浙江省自然科学基金资助项目（397232）
收稿日期：1999-02-28；修订日期：2000-01-25

程起了主要的作用，在破坏含几丁质的中肠围食膜中也可能起重要作用[1]。Hawtin 等[1]通过进行病毒基因缺失研究发现，AcMNPV 无论是缺失了几丁质酶基因还是组织蛋白酶（cathepsin）基因，受其感染的粉纹夜蛾幼虫死亡后尸体会保持数天完整无损。结果表明 AcMNPV 编码的几丁质酶与组织蛋白酶一起，在感染后期的宿主液化过程中起了直接的作用。已发现几丁质酶基因广泛存在于细菌、真菌、线虫、节肢动物、植物等生物体内。但不同几丁质酶基因之间结构和功能差异很大。棉铃虫单粒包埋型核型多角体病毒（HaSNPV）基因组为环状双链 DNA，整个基因组大于 120 kb[2]，根据棉铃虫感染 HaSNPV 死亡后表现的体壁迅速液化症状，HaSNPV 亦应存在编码几丁质酶的基因。为了探索开发利用 HaSNPV 几丁质酶基因以及为今后进行 HaSNPV 重组提供基础，我们克隆了 HaSNPV 几丁质酶基因，并对已报道的杆状病毒的几丁质酶基因进行了进化分析。

1　材料与方法

1.1　材料

HaSNPV 为湖北株，病毒株克隆纯化后，再通过感染棉铃虫 3 龄幼虫进一步在虫体内扩增病毒多角体。限制性内切酶、连接酶及低熔点 agarose 胶均为 GIBCO-BRL 公司产品，Taq DNA pol 酶为永德公司产品，载体为 pUC19，宿主菌为 TG1。

1.2　方法

从病毒多角体中抽提基因组 DNA 的方法见参考文献［2］报道，PCR 扩增正向引物为 AAGAATTCATGAATAATTATTGTTTG，反向引物为 AGAAGCTTCTACAATTCG TTATGAATC，扩增条件为 92 ℃ 1 min，55 ℃ 1 min，72 ℃ 1 min，30 次循环。agarose 分离片段、连接转化、质粒制备与纯化、限制性内切酶反应均按 Sambrook 等[3]方法进行。序列测定应用 ABI PRISM™377DNA 测序仪，反应试剂用 BigDyeTerminator Cycle Sequencing Ready Reaction Kit（PERKIN ELMER）。核苷酸序列和氨基酸序列比较应用 PC\GENE 程序。本文分析所用的各种杆状病毒多角体蛋白和几丁质酶基因数据来源见表 1。

表 1　本文分析所用的各种杆状病毒多角体蛋白和几丁质酶基因数据来源（括号内为 GenBank 接受号）

Table 1　Sequence data of baculovirus polyhedrin genes and chitinase genes used in this study and their sources (Genbank accession number were indicated in the parentheses)

病毒名称 Virus	多角体蛋白基因数据 Polyhedrin gene	几丁质酶基因数据 Chitinase gene
HaSNPV	王　根，张传溪等 1997	本文（AF114795）
HzSNPV	Cowan *et al*. 1997（U67264）	Le *et al*. 1997（U67265）
AcMNPV	Ayres *et al*. 1994（L22858）	Ayres *et al*. 1994（L22858）
BmNPV	Maeda 1994（L33180）	Maeda 1994（L33180）
OpMNPV	Ahrens *et al*. 1997（U75930）	Ahrens *et al*. 1997（U75930）
HcNPV	Croizier *et al*. 1994（D14573）	Gong *et al*. 1999（AF120926）
LdMNPV	Kozio *et al*. 1998（AF081810）	Kozio *et al*. 1998（3822305）
CfNPV	Jamieson *et al*. 1997（U78194）	Arif *et al*.（U72030）

2　结果与分析

2.1　几丁质酶基因的克隆

比较我们已测定的 HaSNPV 的蛋白激酶基因、多角体蛋白基因、ORF1236 及 DNA 聚合酶基因[4~6]与 HzSNPV 相应的基因，发现两种病毒的核苷酸序列同源性高达 97%以上。据此，我们根据已报道的 HzSNPV 几丁质酶基因编码区两端序列，设计了 PCR 引物，为便于克隆和进一步进行表达研究，还根据 HzSNPV 几丁质酶基因编码区 DNA 序列中不存在 *Eco*RⅠ和 *Hin*dⅢ的位点，在 5′端和 3′端 PCR 引物中分别引进了 *Eco*RⅠ和 *Hin*dⅢ的位点。以提取的 HaSNPV 基因组 DNA 为模板，通过 PCR 扩增出 1.7 kb 片段，并克隆于 pUC19 质粒载体中。

2.2　核苷酸序列特征

用 pUC19 质粒载体上的 1 对通用引物 CGCCAGGGTTTTCCCAGTCACGAC 和 GAGCGGATAACAATTTCACACAGG 及进一步合成的 4 个引物 CGTGGCAAGAAAACAATAAACC、CGTGAATTTGTGTTGACATGG、GGATTGTCACTGTTGTAGTTGAGG 及 CGATATACGTGTTATTGTCG，在自动测序仪上测定了所克隆片段的核苷酸全序列。图 1 为 HaSNPV 几丁质酶基因编码区的酶切图谱和测序策略。图 2 为基因编码区的核苷酸和推导的氨基酸序列，编码区全长1 713 bp，G+C 为801 bp，占 46.8%。可编码 570 个氨基酸残基组成的多肽，预计分子量为 63.6 kD。为便于比较，图 2 还在 HaSNPV 几丁质酶基因编码区的核苷酸上方同时列出了 HzSNPV 几丁质酶基因编码区的核苷酸序列。由图 2 可见，HaSNPV 几丁质酶基因与 HzSNPV 几丁质酶基因的同源性很高，序列中不同的核苷酸包括插入、缺失和突变共有 36 个 bp，核苷酸序列一致性高达 98%。但 HaSNPV 几丁质酶基因与 HzSNPV 的相比，核苷酸序列有二处出现很大的变化，其中一处在图 2 的第 495～633 个核苷酸之间，HaSNPV 与 HzSNPV 的相比出现 1 个 bp 的缺失和 7 个 bp 的插入，并使该区域内出现移码，从而导致该区长达 46 个氨基酸残基区段的变化。另一处明显变化是在图 2 第1 677个核苷酸处，HzSNPV 几丁质酶基因出现 15 个 bp 的插入。

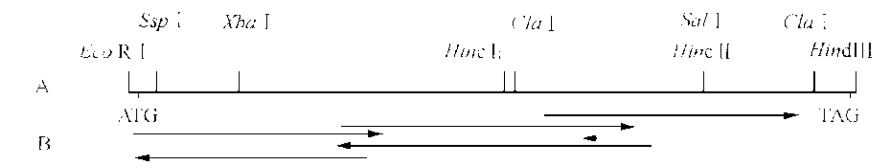


图 1　HaSNPV 几丁质酶基因编码区的酶切图谱（A）和测序策略（B）

Fig. 1　Restriction endonuclease map (A) and the sequencing strategy (B) for the coding region of HaSNPV chitinase gene

2.3　同源性分析

将所推导的 HaSNPV 几丁质酶氨基酸序列与其它已知杆状病毒几丁质酶氨基酸序列进行联配比较，结果列于图 3。在杆状病毒中几丁质酶氨基酸序列很保守，在图 3 的 8 种几丁质酶氨基酸序列中完全相同的氨基酸残基有 245 个，占 42.1%；性质相似的残基有 178 个，占 30.6%；完全相同和性质相似的残基共占 72.7%。HaSNPV 与 HzSNPV 的氨基酸序列非常相

```
....................................................................................
ATGAATAATTATTGTTTGTATTTGTTTGCTTTTAGTATTTTAATATTACATAGCTACGCTGCTCCACCCGGCGTTCCCACATTA   84
 M  N  N  Y  C  L  Y  L  F  A  F  S  I  L  I  L  H  S  Y  A  A  P  P  G  V  P  T  L     28
...................................................T................................
GACTGGGCCGATCACAGTTACGCTTTAGTGCAAATAAATCATCAAGCCACCGCTTACGAATCATTAGTAAACTCTCAACCGTTT  168
 D  W  A  D  H  S  Y  A  L  V  Q  I  N  H  Q  A  T  A  Y  E  S  L  V  N  S  Q  P  F     56
....................................................................................
GTAACTATCAAAGTATCATGGAGCGTATGGTCGGGCGAACAAGGCGAAATCGCTCAAGTGTATCTAGAAAATAAACTAGACTGC  252
 V  T  I  K  V  S  W  S  V  W  S  G  E  Q  G  E  I  A  Q  V  Y  L  E  N  K  L  D  C     84
...................................................................................C
AAAGTAATCAACGTTTGGACTGGTCCGACGCAAGATCGTTTTGCTACTTTTGATTACAATACGAGCGGTCGTTATTCGATGTAT  336
 K  V  I  N  V  W  T  G  P  T  Q  D  R  F  A  T  F  D  Y  N  T  S  G  R  Y  S  M  Y    112
....................................................................................
GTAAAACTGTGCAATGCCGACGGCTGTTCGGCTAGTCAAACCGTAGAAGTCGTGATCGCAGACACCGATGGTAAACATTTAAAA  420
 V  K  L  C  N  A  D  G  C  S  A  S  Q  T  V  E  V  V  I  A  D  T  D  G  K  H  L  K    140
...........................................................................T.........
CCGTTACAATACACGTGGCAAGAAAACAATAAACCGTATACGTACAACACTGATCACACCGTAGCGGCCTATTTT-GTCGAATGG 504
 P  L  Q  Y  T  W  Q  E  N  N  K  P  Y  T  Y  N  T  D  H  T  V  A  A  Y  F  V  E  W    168
--.........--..............-........................................................
GGCGTTTACGGTCGATCTTTTCCCGTGGACAAAGTGCCCACGCCGAACCTTTCGCACATTTTATACGGATTTATACCGATTTGC  588
 G  V  Y  G  R  S  F  P  V  D  K  V  P  T  P  N  L  S  H  I  L  Y  G  F  I  P  I  C    196
...........-...................................-....................................
GGCGGTGACGGTATAAACGACAGTTTAAAATCTATCACGGGGAGTTTTGAAGCGTTGCAAAGATCGTGCGCTGGCAGGGACAAT  672
 G  G  D  G  I  N  D  S  L  K  S  I  T  G  S  F  E  A  L  Q  R  S  C  A  G  R  D  N    224
....................................................................................
TTTAAAGTTTCCATACACGATCCGTGGGCGGCGCTCCAAAAACCACAAACAGGCGTTAGCGCATGGAACGAACCCTACAAAGGC  756
 F  K  V  S  I  H  D  P  W  A  A  L  Q  K  P  Q  T  G  V  S  A  W  N  E  P  Y  K  G    252
....................................................................................
AATTTTGGTCAATTAATGGCAGCAAAATTGGCCAATCCCAATCTAAAAGTGTTGGCATCAATCGGCGGTTGGACACTGTCCGAT  840
 N  F  G  Q  L  M  A  A  K  L  A  N  P  N  L  K  V  L  A  S  I  G  G  W  T  L  S  D    280
....................................................................................
CCCTTCTATCATATGCACGATGCGCGAACGCGACAAATTTTTGTCGAATCCGTGCGTGAATTTGTGTTGACATGGAAATTTTTC  924
 P  F  Y  H  M  H  D  A  R  T  R  Q  I  F  V  E  S  V  R  E  F  V  L  T  W  K  F  F    308
..............................................................C.....................
GACGGCATCGATATTGATTGGGAATTTCCGGGCGGCAAAGGCGCCAATCCAAACGTTGGCGATGTGGAACGCGACAATAACACG 1008
 D  G  I  D  I  D  W  E  F  P  G  G  K  G  A  N  P  N  V  G  D  V  E  R  D  N  N  T    336
....................................................................................
TATATCGCTTTATTGGGCGAATTGCGCGCCATGCTCGATCAAGTTCAAATACAAACGAATCGTACTTTAGAACTCACCACAGCG 1092
 Y  I  A  L  L  G  E  L  R  A  M  L  D  Q  V  Q  I  Q  T  N  R  T  L  E  L  T  T  A    364
....................................................................................
ATTAGCGCTGGCATAGACAAGATCGCCGCGGTTAATTACGACCGAGCACAACAGTATTTAGATAAAATTTTCGTGATGAGTTAT 1176
 I  S  A  G  I  D  K  I  A  A  V  N  Y  D  R  A  Q  Q  Y  L  D  K  I  F  V  M  S  Y    392
....................................................................................
GATTTCAAAGGAGCTTGGTCTAATACCGATCTCGGTCATCAGACGGCACTGTACGGTTCCGCGTGGAAACCTAACGAACCGTAC 1260
 D  F  K  G  A  W  S  N  T  D  L  G  H  Q  T  A  L  Y  G  S  A  W  K  P  N  E  P  Y    420
....................................................................................
ACCGCCAACGTGGCCGTGGACGCTTTGCTCGCGCAACGAGTGAATCCAAAAAAACTCGTGTTGGGCGTAGCAATGTACGGCCGC 1344
 T  A  N  V  A  V  D  A  L  L  A  Q  R  V  N  P  K  K  L  V  L  G  V  A  M  Y  G  R    448
...T.......A........T...G...........................................................
GGTCGGACGGGCGTCCACAACTACAACAGTGACAATCCATTCAGTGGCGTAGCCGTCGGACCGATCACGGGTACATGGGAAAAC 1428
 G  R  T  G  V  H  N  Y  N  S  D  N  P  F  S  G  V  A  V  G  P  I  T  G  T  W  E  N    476
...G................................................................................
GGCATGGTCGACTACAGACAGATTGCTAAAAACATTTCGCGTTACGAATACGCATTCGACGACGTGGCCAAGGCTGCGTACGTA 1512
 G  M  V  D  Y  R  Q  I  A  K  N  I  S  R  Y  E  Y  A  F  D  D  V  A  K  A  A  Y  V    504
................................C...........................................A.......
TTCGACCGCGCCTCTGGCGACTTAATATCGTATGACAGCGAAAGATCGGTACTGGCCAAAGGCGAATACGTGCTTCAGCGTAGA 1596
 F  D  R  A  S  G  D  L  I  S  Y  D  S  E  R  S  V  L  A  K  G  E  Y  V  L  Q  R  R    532
.............................................................A.G.................AAT
CTCGGCGGATTATTCGCATGGGAAATCGATGCGGACAACGGCGATTTGCTCAACGCTATGCGTATTGGCTTGATGAAAACA--- 1677
 L  G  G  L  F  A  W  E  I  D  A  D  N  G  D  L  L  N  A  M  R  I  G  L  M  K  T  -    559
CGCAACGGCCTA.................................... 1722
------------GGTACAAACAGTAGATTGATTCATAACGAATTGTAG 1713
 -  -  -  -  G  T  N  S  R  L  I  H  N  E  L  -  570
```

图2 HaSNPV几丁质酶基因的核苷酸序列和推导的氨基酸序列

Fig. 2 Nucleotide sequence and deduced amino acid sequence of HaSNPV chitinase gene

中间一行为HaSNPV几丁质酶基因核苷酸，在其下一行为推导的氨基酸序列，在其上一行为HzSNPV几丁质酶基因的核苷酸序列。HzSNPV与HaSNPV相同的核苷酸用“·”表示，插入或缺失的核苷酸用“－”表示。图1中标出的限制性内切酶酶切位点用斜体字母表示 The nucleotide sequence of HaSNPV chitinase gene is in the middle line, the deduced amino acid sequence is shown below it, and the nucleotide sequence of HzSNPV chitinase gene is shown above it. The nucleotides in HzSNPV identical to those of HaSNPV are denoted with “·”, and the insertion or gap is denoted with “－”. The restriction enzyme sites indicated in Fig. 1 are in italics

似，相同性高达 90.7%，与 AcMNPV、BmNPV、HcNPV、LdMNPV、OpMNPV 和 CfNPV 的氨基酸序列相同性分别为 64.4%、64.9%、64.2%、62.9%、66.2%和 61.5%。与一种昆虫病原细菌即粘质沙雷氏菌 *Serratia marcescens* 几丁质酶 A 基因（*chi*A）和交替单胞菌 *Alteromonas* sp. 的 *chi*A 也有很高的同源性，分别达到 57.2%和 48.2%。几丁质酶 18 家族的活性位点模式 [LIVMFY] [DN] G [LIVMF] [DN] [LIVMF] [DN] E，其中 G 和 E 两个残基是不变的。在 HaSNPV 几丁质酶中这个活性位点为 FDGIDIDWE（图 3 中第 308～316 个残基），由图 3 可见这在其它杆状病毒中也同样十分保守。其中第 316 位的谷氨酸残基 E 经晶体结构分析被认为可能是几丁质酶的活性位点的催化残基[7]。

2.4　几丁质酶基因的分子进化

图 4（A）是根据已知杆状病毒几丁质酶的氨基酸序列用 PC\GENE 程序绘制的系统树。由图 4（A）可见 HaSNPV 与 HzSNPV 的几丁质酶基因亲缘关系最为接近，两者在系统树中处于一个单独的分枝；而其它 6 种杆状病毒几丁质酶则处于另一分枝中，其中 AcMNPV 与 BmNPV 几丁质酶同源性最高，在 552 个氨基酸序列中有 518 个是一致的，同源性达 93.8%。将杆状病毒几丁质酶系统树与杆状病毒中最为保守的多角体蛋白基因系统树作比较（图 4：B），可见 LdMNPV 与其它杆状病毒亲缘关系较远，在系统树中处于一个独立的分枝，HaSNPV 与 HzSNPV 的多角体蛋白亲缘关系仍是最为接近，在系统树中处于一个分枝；AcMNPV 与 BmNPV 的多角体蛋白进化距离则较两者的几丁质酶进化距离相对为远。这表明不同的基因变异速率是不同的。Hawtin 等[8]曾绘制了 AcMNPV、细菌、真菌和动物的几丁质酶基因系统树，发现 AcMNPV 几丁质酶基因与 *Serratia marcescens* 和 *Alteromonas* sp. 的几丁质酶 A（*chi*A）基因属于同一分枝。我们将 8 种杆状病毒几丁质酶基因和 17 种细菌、真菌、植物和动物的几丁质酶基因进行分子进化分析，8 种杆状病毒几丁质酶基因位于一个系统树小分枝中，杆状病毒几丁质酶基因与 *Serratia marcescens* 和 *Alteromonas* sp. 的几丁质酶 A（*chi*A）基因又属于同一较大分枝中（图未列出），这与 Hawtin 等结果一致。

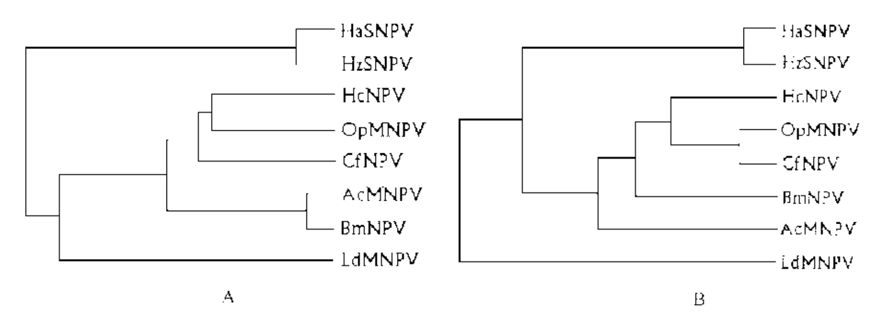


图 4　杆状病毒几丁质酶和多角体蛋白的分子系统树比较

Fig. 4　Comparison of baculovirus chitinase and polyhedrin phylogenetic trees

A. 几丁质酶系统树 Chitinase phylogenetic tree；B. 多角体蛋白系统树 Polyhedrin phylogentic tree

3　讨论

通过 PSORT Ⅱ server 中 von Heijne 的信号序列识别程序计算，HaSNPV 几丁质酶 N 端可

```
HaSNPV  MNNYCLYLFAFSILILHSYAAPPGVPTLDWADHSYALVQINHQATAYESLVNSQPFVTIKVSWSVWSGEQGEIAQ  75
HzSNPV  MNNYCLYLFAFSILILHSYAAPPGVPTLDWADHSYALVQINHQATSYESLVNSQPFVTIKVSWSVWSGEQGEIAQ  75
HcNPV   MLNYLVSILWVTAALSY---ASPGTPVIDWADRNYALVKINADATAYENLIERKEQVNVQVSWNVWNGDVGDMAY  72
OpMNPV  MLHYLATILWL--AVAH---ASPGTPVIDWADRNYALVSVNSEATAYENLVERKAGVSVPVSWNVWNGGVGDMAY  70
CfNPV   MVHYLVKVLWLTIAVAY---ASPGTPVIDWADRNYALVKINSDATAYENLIQRENHVNVQVSWNVWNGGVGDMAY  72
AcMNPV  MLYKLLNVLWLV-AVSN---AIPGTPVIDWADRNYALVEINYEATAYENLIKPKEQVDVQVSWNVWNGDIGDIAY  71
BmNPV   MLYKLLNVLWLVVAVSN---AIPGTPVIDWAERNYALVKINYEATAYENLIKLKEQVDVHVSWNVWNGDIGDIAY  72
LdMNPV  MIRYIAFVLAAASIATA---APPGAPVIDWADRNYALVEVSPDAVAYERLVKRKDRVEVNVSWSVYTGEPGDTAY  72
        *      ..            * **.*..***...**** .. .*..**.*. .   *..  ***.*..*. *. *

HaSNPV  VYLENKLDCKVINVWTGPTQDRFATFDYNTSGRYSMYVKLCNADGCSASQTVEVVIADTDGKHLKPLQY-TWQEN  149
HzSNPV  VYLENKLDSKVINVWTGPTQDRFATFDYNTSGRYSMYVKLCNADGCSASQTVEVVIADTDGKHLKPLQY-TWQEN  149
HcNPV   VLFDDK-------QVWKGYAAAKKASVNVVKSGQFGMRVKLCDDDGCSVSEPVTVKVADTDGGHLLPMEY-VWQEN  140
OpMNPV  VLFNEN-------QVWKGAAAAKRATIDVSKSGQFNMRVKLCDDDGFSVSEPVTVRVADTDGGHLSPLEY-AWGEN  138
CfNPV   VLFNDK-------QVWKGDADAKKATVVVSKSGQFRMHVKVCDEDGCSVSEPVTVKVADTDGGHLNPLEY-VWREN  140
AcMNPV  VLFDEQ-------QVWKGDAESKRATIKVLVSGQFNMRVKLCNEDGCSVSDPVLVKVADTDGGHLAPLEY-TWLEN  139
BmNPV   VFFDEQ-------QVWKGDADSKRATINVIVSGQFNMRVKLCNEDGCSISDPVLVKIADTDGGHLAPLEY-TWLEN  140
LdMNPV  VLFGDR-------QVWVGEAAGKRASVCVSSGGKHEMRVKLCNADGCSVSAPVTVKIADTDGAHLDPLPYDAWLEN  141
        * ....         .** * .... *.      .*.   * **.*..** * *..* * .***** ** *.  * .* **

HaSNPV  NKPYTYNTDHTVAAYFVEWGVYGRSFPVDKVPTPNLSHI---LYGFIPICGGDGINDSLKSITGSFEALQRSCAG  221
HzSNPV  NKPYTYNTDHTVAAYFC-RMAFTRSFPVTKCPRRTFRTFYTDLYRFAAVTD----KRQFKIYHGEFEALQRSCAG  219
HcNPV   NKP-GRRANKTVAAYFVEWGVYGRNFPIDKVPLPNLSHL---LYGFIPICGGDGLNDALKTIPGSFEALQRSCKG  211
OpMNPV  NKP-GRPHNKTVAAYFVEWGVYGRGFPVDKVPLPNLSHL---LYGFIPICGGDGLNDALKTIPGSFEALQRSCKG  209
CfNPV   NKP-GRRQDKTIAAYFVEWGVYGRNFPVDKVPLPNLSHL---LYGFIPICGGDGLIDALKTIPGSFEALQRSCQG  211
AcMNPV  NKP-GRREDKIVAAYFVEWGVYGRNFPVDKVPLPNLSHL---LYGFIPICGGDGINDALKTIPGSFESLQRSCKG  210
BmNPV   NKP-GRREDKIVAAYFVEWGVYGRSFPVDKVPLPNLSHL---LYGFIPICGGDGINDALKTIPGSFEALQRSCRG  211
LdMNPV  NRP-AEPTGKVVAAYFVEWGVYGRDFPADKVPVPNLTHL---LYGFVPMCGGDGINDSLKSIAGSFEALQRSCAG  212
        *.* .     ....****  . ...*.** .* * ...   .     ** * .. .          .*    *.**.***** *

HaSNPV  RDNFKVSIHDPWAALQKPQTGVSAWNEPYKGNFGQLMAAKLANPNLKVLASIGGWTLSDPFYHMHDARTRQIFVE  296
HzSNPV  RDNFKVSIHDPWAALQKPQTGVSAWNEPYKGNFGQLMAAKLANPNLKVLASIGGWTLSDPFYHMHDARTRQIFVE  294
HcNPV   RADFKVAIHDPWAAIQKPQKGVSAWNEPYKGNFGQLMAAKLANPHIKILPSIGGWTLSDPFYFMHDAENGNVFVD  286
OpMNPV  RADFKVAIHDPWAAIQKPQKGVSAWNEPYKGNFGQLMAAKLANPHLKILPSIGGWTLSDPFYFMHDADKRRVFVE  284
CfNPV   RSDFKVAIHDPWAAIQKPQKGVSAWNEPYKGNFGQLMAAKLANPHLKILPSIGGWTLSDPFYYMHDAIKRRVFVD  286
AcMNPV  REDFKVAIHDPWAAVQKPQKGVSAWNEPYKGNFGQLMAAKLANPHLKILPSIGGWTLSDPFYFMHDVEKRNVFVD  285
BmNPV   REDFKVAIHDPWAAVQKPQKGVSAWNEPYKGNFGQLMAAKLANPHLKILPSIGGWTLSDPFYFMHDVEKRNVFVE  286
LdMNPV  RDDFKLAVHDPWAAVQKPQKGVTAWSEPYKGNFGQLMAIKKAHPHLRVLPSIGGWTLSDPFYHMHDASARRAFVE  287
        *..**...******.****.**.**.************ * *.*....*.************ ***.    .**.

HaSNPV  SVREFVLTWKFFDGIDIDWEFPGGKGANPNVGDVERDNNTYIALLGELRAMLDQVQIQTNRTLELTTAISAGIDK  371
HzSNPV  SVREFVLTWKFFDGIDIDWEFPGGKGANPNVGDVERDNNTYIALLGELRAMLDQVQIQTNRTLELTTAISAGIDK  369
HcNPV   SVKEFLQVWKFFDGVDIDWEFPGGKGANPALGNGERDAHTYIILLQELRSMLDDLQLQTGKTYVLTSAISAGHDK  361
OpMNPV  SVKEFLQVWKFFDGVDIDWEFPGGKGANPALGNGERDADTYLVLLKELRAMLDELQLQTNKTYELTSAISSGYDK  359
CfNPV   SVKEFLQVWKFFDGVDIDWEFPGGKGANPTLGNKERDADTYNILLKELRAMLDELQLQTGKTYELTSAISSGHDK  361
AcMNPV  SVKEFLQVWKFFDGVDIDWEFPGGKGANPSLGDADGDAKTYILLLEELRAMLDDLEAQTGRVYELTSAISAGYDK  360
BmNPV   SVKEFLQVWKFFDGVDVDWEFPGGKGANPSLGDAERDAKTYILLLDELREMLDDLEVQTGRVYELTSAISAGYDK  361
LdMNPV  SAREFLQTWKLFDGLDVDWEFPGGKGANPLLGDAERDRETYALLLEELRLMLDELQAETNRTYELTSAIGAGDDK  362
        *..**. .**.***.*.************ .*.  . * .**  ** *** ***... .*...  **.**..* **

HaSNPV  IAAVNYDRAQQYLDKIFVMSYDFKGAWSNTDLGHQTALYGSAWKPNEPYTANVAVDALLAQRVNPKKLVLGVAMY  446
HzSNPV  IAAVNYDRAQQYLDKIFVMSYDFKGAWSNTDLGHQTALYGSAWKPNEPYTANVAVDALLAQRVNPKKLVLGVAMY  444
HcNPV   IAVVQYEAAQKYLDKIFLMSYDFKGAWSNTDLGYQTTLYAPSWNPDELYTTDHAVKLLTSQGVASDKIIVGIAMY  436
OpMNPV  IAVVKYDAAQRFLDKIFLMSYDFKGAWSNTDLGYQTTLYAPSWNANELYTTDHAVKLLTGQGVAAHKLIVGVAMY  434
CfNPV   IAVVKYDLAQKFLDKIFLMSYDFKGAWSNIDLGYQTTLYAPSWNPNELYNTDYAVKSLTNQGVDPRKIIVGVAMY  436
AcMNPV  IAVVNYAEAQKSLGKIFLMSYDFKGAWSNTDLGYQTTVYAPSWNSEELYTTHYAVDALLKQGVDPNKIIVGVAMY  435
BmNPV   IAVVNYAEAQKSLGKIFLMTYDFKGAWSNTDLGYQTTVYAPSWNSEELYTTHYAVDALLEQGVDPNKIIVGVAMY  436
LdMNPV  IAVVDYGRAHRSLDRIFLMSYDFKGAWSTTELGHQTALYAPSWRPDEPYTADGAVRRLLAQGAPADKIVLGVAMY  437
        **.*.*. *.. *..**.*.********...**.**.*..*...*...* *...  **  *  *  .  *...*.***
```

（下转 239 页）

```
HaSNPV   GRGRTGVHNYNSDNPFSGVAVGPITGTWENGMVDYRQIAKNIS--RYEYAFDDVAKAAYVFDRASGDLISYDSER   519
HzSNPV   GRGWTGVHNYDSDNPFSGVAVGPITGTWENGVVDYRQIAKNIS--RYEYAFDDVAKAAYVFDRASGDLISYDSER   517
HcNPV    GRGWTGVAGYDGDNFFSGTAKGPISGTWENGVVDYRQIKNELN--TYLYRFDTAAKAAYAFNSAKGDLISFDSVD   509
OpMNPV   GRGWTGVSGYAGDKYFSGTADGPVPGTWENGVVDYRQINNELS--KYIYRFDAAAKAAYVFNKERGDLISFDSVD   507
CfNPV    ARGWTGVAGYDGDNYFTGTAAGPVTGTWEDGVVDYRQIKNELN--KYKYKFDTAAKASYVFNKDNGDLISFDSVD   509
AcMNPV   GRGWTGVTNYTNDNYFSGTGNGPGSGTWEDGVVDYRQIQKDLN--NYVYTFDSAAQASYVFDKSKGDLISFDSVD   508
BmNPV    GRGWTGVTNYTNGNYFSGTGNGPVSGTWEDGVVDYRQIQKDLN--NYVYTFDSAAQASYVFDKSKGDLISFDSVD   509
LdMNPV   GRGWTGVTGYADGNPFTGAATGPVPGTWEDGVVDYRELRAAIDGGRYAYAYDSTAKGAYAFDETAGALVSFDSAE   512
         .**.*** .*....  *.*..  ** .****.*.****..   ...    * * .*..*...*.*.  .  *.*.*.**
HaSNPV   SVLAKGEYVLQRRLGGLFAW-EIDA DNGDLLNAMRIGLMKT-----GTNSRLIHNEL   570
HzSNPV   SVLAKGEYVLKRRLGGLFAW-EIDA DNGDLLNAMHVGLMKTNRNGLGTNSRLIHNEL   573
HcNPV    SVLAKSVYVEQNNLGGVFAW-EIDA DNGDLLNAMNQQTKLL--K----------DEL   553
OpMNPV   SVLAKNVYVQQNGLGGLFAW-EIDA DNGDLLNAMNERVRV---K----------DEL   550
CfNPV    SVLAKTSYVDPKWLGWVICLGDPTP TNGDLLNVMNKNFKI---K----------DEL   553
AcMNPV   SVLGKVKYVDRNKLGGLFAW-EIDA DNGDLLNAINAQFK-P--K----------DEL   551
BmNPV    SVLGKVKYVDRNNLGGLFAW-EIDA DNGDLLNAMNAQFK-L--R----------DEL   552
LdMNPV   SILDKGKYARELGLGGLFAW-EIDA DNGDLLNAMHAGLKPAARR----------VEL   558
         .*.*.*  *.      ** ..    . .. .******...   ..          **
```

图 3　8 种杆状病毒几丁质酶氨基酸序列联配

Fig. 3　Alignment of the amino acid sequence of 8 baculovirus chitinases

在 8 种杆状病毒完全保守的氨基酸残基用“*”表示；性质相似的氨基酸残基用“.”表示；

几丁质酶 18 家族的活性位点 FDGVDIDWE 用黑体字母表示

The perfectly conserved amino acids are denoted with asterisks, the amino acids with similar character are indicated with dots. The catalytic domain of chitinase family 18 is in boldface

能有一由 20 个氨基酸残基信号肽序列。粘质沙雷氏菌几丁质酶 A 基因（*chi*A）的 N 端有一个原核生物的分泌信号肽，但 AcMNPV 的几丁质酶没有相同的分泌信号肽，代之的可能有一个真核信号肽[9]，但该信号肽似没有直接起有效的蛋白分泌作用，因为病毒感染的细胞内几丁质酶有 90％以上的活性[8]。HaSNPV 几丁质酶的信号肽有待于进一步的研究。

几丁质为聚-β-1，4-N-乙酰葡糖胺，广泛存在于自然界，节肢动物表皮、真菌细胞壁等均含有几丁质，含量 25％～60％左右，是使表皮呈现弹性、韧性的主要因素。许多生物从细菌到高等植物均能产生降解几丁质的几丁质酶，细菌的几丁质酶使细菌能利用几丁质作为 C 源和 N 源，植物的几丁质酶可抵抗病原物的侵染。昆虫也编码几丁质酶，在蜕皮液中的几丁质酶可在昆虫脱皮时用于降解几丁质的旧表皮。AcMNPV 几丁质酶基因 *chi*A 是一晚期基因，在病毒感染细胞后 8 h 开始表达，12 h 达到高峰，酶活性一直持续到感染末期[8]。昆虫表皮中含有大量的几丁质，被杆状病毒感染的昆虫均最终体壁十分脆弱易破，已明确杆状病毒编码的几丁质酶和在基因组中相邻的半胱氨酸蛋白酶能使昆虫体壁迅速降解液化，促使病毒释放出昆虫体外[1]。一种缺少蛋白酶和几丁质酶的 *Serratia marcescens* 株系较野生的株系对果蝇的毒力要低[10]。将一种非专一性几丁质酶与 NPV 混合使用杀舞毒蛾 2 龄幼虫，比单独用 NPV 的杀虫速度要快得多[11]。这两研项究结果表明几丁质酶可能破坏含几丁质的围食膜，从而使多角体中释放出的病毒粒子能更有效地到达中肠表皮细胞。AcMNPV 的多角体中也检测到了有活性的几丁质酶，而且病毒编码的几丁质酶在 pH3.0～10.0 均有活性，而 *Serratia*

marcescens 编码的几丁质酶虽然与NPV编码的几丁质酶有高达57%的序列相同性，但其则在pH>7时就会很快失去活性[1]。这提示病毒编码的几丁质酶可能在幼虫中肠液碱性环境中仍能对含几丁质的围食膜起分解作用。

AcMNPV的几丁质酶与 *Serratia marcescens* 的*chi*A序列有60.5%的相同性，而且AcMNPV的*chi*A基因核苷酸序列中G+C含量占50.4%，明显高于其基因组G+C占41%的比例，而与 *Serratia marcescens* 的*chi*A高G+C（58.6%）偏爱相类似，因而推测AcMNPV或杆状病毒祖先是经水平转移从细菌获得*chi*A[8]。HaSNPV几丁质酶核苷酸序列与HzSNPV几丁质酶核苷酸序列一致性高达98%，其中G+C比例都为46.8%，相对HzSNPV多角体蛋白基因区附近已测的14 235对核苷酸中，G+C只占36.2%来看，几丁质酶基因G+C比例高出达10.6% 表明杆状病毒几丁质酶基因核苷酸序列中G+C含量的确相对较高，这是由于杆状病毒的基因来源于细菌之故还是由于几丁质酶存在其它选择压力尚不清楚。我们分析表明杆状病毒几丁质酶基因与已报道的烟草天蛾 *Manduca sexta* 几丁质酶基因则几乎没有同源性（资料未列）。

根据与其它杆状病毒几丁质酶氨基酸序列进行联配比较的分析结果（图3），HaSNPV几丁质酶核苷酸序列中第495～633个核苷酸所编码的46个氨基酸序列，与AcMNPV、BmNPV、CfMNPV、HcNPV、OpMNPV和LdMNPV的几丁质酶相应氨基酸很相似，其中有32个氨基酸残基在7种病毒中完全一样，而HzSNPV几丁质酶的该区域氨基酸序列与其它杆状病毒相比则有很大的不同，显然HaSNPV几丁质酶较HzSNPV更为保守。但HaSNPV和HzSNPV几丁质酶基因在该区域的核苷酸序列仍是十分相同（图2），这可能表明HzSNPV几丁质酶基因是从HaSNPV几丁质酶基因经缺失和插入突变而来的。

本文结果表明杆状病毒几丁质酶基因系统树与多角体蛋白基因系统树有较大的不同，如BmNPV与AcMNPV的多角体蛋白进化距离相对较远，而两者的几丁质酶亲缘关系则非常近，而根据已测定的AcMNPV、BmNPV、OpMNPV和LdMNPV基因组全序列看，AcMNPV与BmNPV两种病毒基因组无论是各基因的线性排列还是核苷酸序列都很相似，两者的亲缘关系无疑是很接近的，OpMNPV与AcMNPV和BmNPV基因组无论在各基因的线性排列还是核苷酸序列方面都有相对较大的距离，而LdMNPV基因组与上述三种病毒的差异则更大。根据不同基因序列比较绘制的系统树不尽相同，这一方面说明了不同来源和不同功能的基因进化速率是不一样的；另一方面也说明了杆状病毒每一个单独的基因进化情况虽能在一定程度上反映杆状病毒的分子系统发育，但并不能代表病毒整体的分子系统发育情况，只有通过多基因甚至整个基因组的进化分析，才能准确确立整体的杆状病毒的分子系统树。

参 考 文 献（References）

[1] Hawtin R E, Zarkowwska T, Arnold K *et al*. Liquefaction of *Autographa californica* nucleopolyhedrovirus-infected insects is dependent on the integrity of virus-encoded chitinase and cathepsin genes. Virology, 1997, 238: 243～253

[2] 王 根，张传溪，金 伟等. 棉铃虫核型多角体病毒限制性酶切图谱和多角体蛋白基因定位. 环境与应用生物学报，1996，2（4）：387～394

[3] Sambrook J, Fritsch E F, Maniatis T. Molecular Cloning: A Laboratory Manual (2nd ed). Cold Spring Harbor Laboratory Press, 1989

[4] 王　根，张传溪，贡成良等. 棉铃虫核型多角体病毒多角体蛋白基因的克隆和核苷酸序列分析. 病毒学报，1997，13（1）：82～86

[5] 王　根，张传溪，金　伟等. 棉铃虫核多角体病毒DNA聚合酶基因的克隆和核苷酸序列分析. 中国病毒学，1998，13（1）：77～82

[6] 张传溪，王　根，胡　萃等. HaSNPV蛋白激酶基因的核苷酸序列分析. 生物物理与生物化学学报，1997，29（3）：322～326

[7] Perrakis A，Tews I，Dauter Z *et al*. Crystal structure of bacterial chitinase at 2.3 resolution. Structure，1994，2：1 169～1 180

[8] Hawtin R E，Arnold K，Ayres M D *et al*. Identification and preliminary characterization of a chitinase gene in the *Autograph californica* nuclear polyhedrosis virus genome. Virology，1995，212：673～685

[9] Ayres M D，Howard S C ，Kuzio J *et al*. The complete DNA sequence of *Autographa californica* nuclear polyhedrosis virus. Virology，1994，202：586～605

[10] Flyg C，Boman H G. *Drosophila* genes cut and miniature are associated with the susceptibility to infection by *Serratia marcescens*. Genet. Res.，1988，52：51～56

[11] Shapiro M，Preisler H K，Robertson J L. Enhancement of baculovirus activity on gypsy moth（Lepidoptera：Lymantriidae）by chitinase. J. Econ. Entomol.，1987，80：1 113～1 116

Chitinase gene of *Helicoverpa armigera* nucleopolyhedrovirus and the phylogenetic tree of the baculovirus chitinase genes*

ZHANG Chuan-xi[1]，LIN Xin-da[1]，WU Jun[2]

（1. Institute of Applied Entomology，Zhejiang University，Hangzhou　310029，China；
2. Shanghai Institute of Biochemistry，the Chinese Academy of Sciences，Shanghai　200031，China）

Abstract：The chitinase gene of *Helicoverpa armigera* single enveloped nucleocapsid nucleopolyhedrovirus（HaSNPV）was amplified by using PCR method. The nucleotide sequence of the coding region was determined. The total length of the coding region was 1 713 bp and was potential to encode a 570 amino acid polypeptide with predicted molecular weight of 63.6 kD. The alignment of HaSNPV chitinase amino acid sequence with those of other 7 reported baculovirus chitinases showed that HaSNPV chitinase was most closely related to HzSNPV chitinase，with 90.7% amino acid identity. It also shared a high degree of homology with chitinases of AcMNPV，BmNPV，HcNPV，LdMNPV，OpMNPV and CfMNPV，with 64.4%，64.9%、64.2%、62.9%、66.2% and 61.5% amino acid identity，respectively. A phylogenetic tree of 8 baculovirus chitinases was drawn by using PC\GENE program，and compared with the phylogenetic tree based on the polyhedrin amino acid sequences of the same 8 baculoviruses. The result showed that the evolutionary rate of chitinase genes was in some degree different to that of polyhedrin genes.

Key words：*Helicoverpa armigera* nucleopolyhedrovirus；chitinase gene；nucleotide sequence；baculovirus；phylogenetic tree

* The nucleotide sequence data reported here have been deposited with the GenBank database and assigned the accession number AF114795